Bioorganic & Medicinal Chemistry Letters 23 (2013) 6667–6672

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Comprehending renin inhibitor's binding affinity using structure-based approaches

Govindan Subramanian [a,*], Shashidhar N. Rao [b,†]

[a] VMRD Global Discovery, Zoetis, 333 Portage Street, Kalamazoo, MI 49007, USA
[b] Department of Chemistry and Chemical Biology, Rutgers University, 610 Taylor Road, Piscataway, NJ 08854, USA

ARTICLE INFO

Article history:
Received 7 September 2013
Revised 19 October 2013
Accepted 22 October 2013
Available online 31 October 2013

Keywords:
Renin inhibitor
MM-GB/SA
3D-QSAR
In silico modeling
MD simulation

ABSTRACT

The performance of several structure-based design (SBD) approaches in predicting the binding affinity of diverse small molecule inhibitors co-crystallized to human renin was assessed to ascertain the modeling tool and method of choice required when dealing with structure-based lead optimization projects. Most of the SBD approaches investigated here were able to provide qualitative guidance, but quantitative accuracy as well as decisive discrimination between [in]actives is still not within reach. Such an outcome suggests that the current methods need improvement to capture the overall physics of the binding phenomenon for consistent applications in a lead optimization setting.

Predicting small molecule binding affinity when bound to its intended target is still a challenging but formidable problem in in silico computer-assisted molecular discovery (CAMD) ecosystem. A number of commercial vendor and academic offerings to model affinity using structure-guided and ligand-based approaches have not yet offered unequivocal solutions that could be applied a priori in a reliable and consistent manner. Consequently, the field can be considered to still be in its infancy with several publications showcasing their attempts to model affinity using analogous compound datasets. These range from the use of simple atomic and molecular descriptors of small molecules to highly sophisticated molecular dynamics (MD) simulations using ligand bound protein structures that attempts to quantify affinity and allows comparison to experimental values.

The choice of the methods and the application tools mostly depend on the information at hand. 3-Dimensional quantitative structure activity relationship (3D-QSAR) techniques like CoMFA®,[1] CoMSIA[2] (using molecular superpositions) and CATALYST[3] (using pharmacophore features) have been the hallmark for affinity predictions for the past several decades. Recent advances in the X-ray structure determination for a large set of small molecule bound to proteins as well as improvements in silico computational docking methods enabled the application of linear interaction energy (LIE)[4] and molecular mechanics with generalized Born surface area (MM-GB/SA)[5] or Poisson–Boltzmann (MM-PB/SA)[6] to prevail for predicting relative binding affinity. Rigorous free energy perturbation (FEP),[7] thermodynamic integration (TI),[8] etc., methods involving long time MD simulations is not yet practical for evaluating large datasets or modeling analogs with larger modifications.

Although a general consensus has not emerged from the application of these structure-based tools, their increasing prominence in literature can be appreciated from some of the recent publications.[9] For instance, Srivastava and Sastry[10] reported the results of MM-GB/SA and MM-PB/SA computed free energies from MD simulation trajectories and showed the sensitivity of the predicted activity to be dependent on the simulation time, while Genheden and Ryde provide hints[11] to achieve statistical validation in predicting binding energies via MD simulations. In addition, the latter authors suggest that MD simulations on regions proximal to the binding site can improve the efficiency of the free-energy calculations.[12] In an apparent trade-off for simulating long time and treating the solvent explicitly, continuum solvation-based methods (solvent treated implicitly) offer a promising intermediate solution and have gained widespread popularity for its speed and ability to predict affinity trends reliably.[13] Representative among them are the work from Kohlmann et al.[14] and Rapp et al.[15] who report a good correlation between predicted (MM-GB/SA) and experimental observations (IC_{50}) for congeneric small molecule chemical series. Recently, Hotiana and Haider[16] used MM-GB/SA approach to understand and predict the resistance mutations in the HCV NS3/4A serine protease binding site while Ryde and coworkers compared the performance of several structure-based techniques in

* Corresponding author. Tel.: +1 269 833 2904; fax: +1 269 833 4627.
E-mail address: govindan.subramanian@zoetis.com (G. Subramanian).
† Current address: AMRI Global, 26 Corporate Circle, Albany, NY 12203, USA.

0960-894X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmcl.2013.10.044

predicting the relative affinity trends.[17] Finally, Greenidge et al.[18] computed the MM-GB/SA free energies for a diverse set of 855 protein–ligand complexes and compared against the experimental data (K_i/K_d) to reveal good correlations ($r^2 \sim 0.63$), suggesting the computed free energy indeed captures several aspects of ligand binding.

In spite of these well investigated studies, a systematic effort has not been undertaken to demonstrate if the additional computational investment is beneficial and to what extent the predictions improve the decision making process at the lead optimization phase of the project. Similarly, the effect or efficiency of SBD approaches in predicting binding affinity for a single target with diverse chemical matter (not congeners) and adaptive protein structures (induced fit) has not been elaborately examined. We attempt to address the above using human renin as the protein target and the small molecule or peptidomimetic inhibitors bound crystallographically to renin.

Renin is a therapeutically validated aspartyl protease in the renin–angiotensin–aldosterone (RAAS) system with aliskiren being the first FDA approved direct renin inhibitor drug to treat hypertension.[19] The low oral bioavailability of aliskiren prompted other pharmaceutical companies to pursue this target with only two clinical candidates (ACT-077825[20] and VTP-27999[21]) reported till date. This demonstrates the difficulty in identifying small molecules that could be drugged and hence there is immense interest in this area.

54 small molecule and peptidomimetics bound renin X-ray co-complexes were retrieved from the PDB (Protein Data Bank: http://www.rcsb.org). The experimental IC_{50} values reported in the source literature were used in 46/54 cases with 28 inhibitors exhibiting $IC_{50} \leqslant 10$ nM. Publically available experimental data could not be obtained for the remaining eight small molecule bound renin structures. The physicochemical property profiles of all the 54 compounds reveal that the molecular weight values range between 293 and 730; H-bond acceptor and donor range between 2 and 13 and 2–7 respectively with 42/54 inhibitors failing the Lipinski's filter[22] in spite of exhibiting favorable ligand efficiency index[23] (Table S1 in supporting information). The high flexibility of the inhibitors is also evident from the variations (2–24) in the number of rotatable bonds.

Preliminary analysis of the small molecule renin inhibitors showed that the bound inhibitors are structurally diverse (Table S2 in supporting information) and also occupy different binding subsites when compared to the bound peptidomimetic, Aliskiren (PDB: 2V0Z). This clearly emphasizes that standard ligand-based molecular overlay techniques are not optimal to develop quantitative models and that structure-based techniques are the approach of choice.

Each protein–ligand complex was prepared using the 'protein preparation' wizard within MAESTRO (v9.3.5) software suite from Schrodinger Inc., with the default setting. All the solvent waters were removed to maintain uniformity in the starting structures. In cases where multiple chains were reported, only one of the protein chains was considered. Each of the initial complexes was pairwise structurally aligned to the prototype 2V0Z (PDB code) renin-aliskiren complex using the 'structalign' script within the same package. The nature of the binding site was assessed using the SiteMap module within MAESTRO and the results reveal the Dscore score to be ~1.112 on average (Table S3 in supporting information) reflecting the *druggability* of the site.[24] As a first step, the complexes were 'scored in place' and the corresponding Glide scores from standard precision (SP) and extra precision (XP) modes saved. The MM-GB/SA and linear interaction approximation (LIA)[25] approaches implemented within MAESTRO were used to compute the binding free energy of each complex independently. The complexes were also evaluated additionally via the MM-GB/SA and MM-PB/SA approaches implemented within AMBER (v12.0) software modules. Finally each complex was solvated with a truncated octahedral box of TIP3P water molecules and subjected to MD simulations for 2 ns following a 0.5 ns equilibriation of the solvated system. The post-equilibrium trajectories were analyzed at each 1 fs interval to compute the ligand binding affinity using the 'MMPBSA.py' python script in AMBER. In these analyses, the MM-GB/SA binding affinities were evaluated after the water molecules and counterions solvating the systems were removed.

To be as close as possible to the perceived reality, all the complexes were treated independently and the overall structural arrangement observed in the X-ray structure was kept as such. This auger well with the recent study that showed minor perturbations to the ligand in the binding site can result in dramatic changes of the computed MM-GB/SA energies.[26] Consequently, no cross-docking or the effect of alternate binding site 'loop conformation' was considered explicitly. The only approximation relates to the removal of explicit solvent water molecules in the investigated complexes, even if they were supposed to make solvent mediated hydrogen bonds to stabilize the protein–ligand complex.

Table 1 summarizes the experimental IC_{50} results (where available) as well as prediction from the above computational approaches. Figure 1 displays the correlation/scatter plots on the comparisons between experiment and prediction from different computational modeling approaches. Since experimental measurement error(s) were not provided in the original literature for majority of the renin inhibitors and the fact that control/tool compounds were not reported as well (to ascertain inter laboratory measurement variability), all the experimental results comparison with prediction should be treated appropriately. It should also be emphasized that experimental IC_{50} values (not K_i/K_d) are compared against the computed binding energies/affinities and docking scores.

As seen in Figure 1, the initial assessment of the outcome from multiple SBD approaches pursued in this study is not striking, but certainly very encouraging. Our analyses of the X-ray poses find that MAESTRO MM-GB/SA binding energies seem to correlate with pIC_{50} values with better statistics ($r^2 \sim 0.52$) compared to the corresponding AMBERS MM-GB/SA binding energies ($r^2 \sim 0.26$). This difference may be attributed to several factors including protein preparation, force–field parameters, assigned partial charges, treatment of hydrophobic effects and inter-residue electrostatic interactions between MAESTRO and AMBER. However, it is noteworthy to observe that the computed ΔG_{bind} between MAESTRO and AMBER implementations of MM-GB/SA are linearly correlated with an $r^2 \sim 0.74$ (Fig. 1d). Interestingly, the computed ΔG_{bind} energy and its spread are much tighter in AMBER but wider in MAESTRO (Fig. 1a, b). The performance outcome from MM-PB/SA was significantly weaker (Fig. 1c) compared to its corresponding MM-GB/SA within AMBER. A similar trend between the MM-GB/SA and MM-PB/SA approaches was reported by Hou et al.[27] from their investigation of 98 protein–ligand crystal structure complexes. We believe that the results from MM-PB/SA could be improved by including explicit solvent waters that may bridge the protein–ligand interaction as has been suggested recently,[28] but is beyond the scope of this work. The results from Liaison (LIA implementation within MAESTRO) seem to be intermediary in performance (Fig. 1f), in spite of linearly regressing the computed van der Waals, electrostatic, and cavity feature terms to experimental affinity. We anticipate that additional system-dependent descriptions may allow for a better correlation as has been reported recently.[29] Such customizations and selected inferences from literature is likely to improve the statistical outcome, but is highly dependent on the system of interest. As more terms are added to the regression, there is a good probability for a better statistical fit, but such a process diminishes